UNCECOMP 2015
1st ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Crete Island, Greece, 25–27 May 2015

STOCHASTIC REPRESENTATIONS AND STATISTICAL INVERSE IDENTIFICATION FOR UNCERTAINTY QUANTIFICATION IN COMPUTATIONAL MECHANICS

C. Soize[1], C. Desceliers[1], J. Guilleminot[1], T.T. Le[1], M.T. Nguyen[1], G. Perrin [1,2,3], J.M. Allain[2], H. Gharbi[2], D. Duhamel[3], and C. Funfschilling[4]

[1]Université Paris-Est
Modélisation et Simulation Multi Echelle, MSME UMR 8208 CNRS
5 bd Descartes, 77454 Marne-la-Vallée, France
e-mail: christian.soize@univ-paris-est.fr

[2] Ecole Polytechnique
Laboratoire de Mécanique des Solides, 91128, Palaiseau cedex, France

[3] Université Paris-Est
Laboratoire Navier (UMR 8205), CNRS, ENPC, IFSTTAR, F-77455 Marne-la-Valle, France

[4] SNCF, Innovation and Research Department
Immeuble Lumière, 40 avenue des Terroirs de France, 75611 Paris, France

Keywords: Uncertainty quantification, Statistical inverse problems, Stochastic modeling, Random fields, Polynomial Chaos, Multiscale analysis.

Abstract. *The paper deals with the statistical inverse problem for the identification of a non-Gaussian tensor-valued random field in high stochastic dimension. Such a random field can represent the parameter of a boundary value problem (BVP). The available experimental data, which correspond to observations, can be partial and limited. A general methodology and some algorithms are presented including some adapted stochastic representations for the non-Gaussian tensor-valued random fields and some ensembles of prior algebraic stochastic models for such random fields and the corresponding generators. Three illustrations are presented: (i) the stochastic modeling and the identification of track irregularities for dynamics of high-speed trains, (ii) a stochastic continuum modeling of random interphases from atomistic simulations for a polymer nanocomposite, and (iii) a multiscale experimental identification of the stochastic model of a heterogeneous random medium at mesoscale for mechanical characterization of a human cortical bone.*

1 INTRODUCTION

The problem related to the identification of a model parameter (scalar, vector, field) of a boundary value problem (BVP) (for instance, the coefficients of a partial differential equation) using experimental data related to a model observation (scalar, vector, field) of this BVP, is a problem for which there exists a rich literature, including numerous textbooks. In general and in the deterministic context, there is not a unique solution because the function, which maps the model parameter (that belongs to an admissible set) to the model observation (that belongs to another admissible set) is not a one-to-one mapping, and consequently, cannot be inverted. It is an ill-posed problem. However, such a problem can be reformulated in terms of an optimization problem consisting in calculating an optimal value of the model parameter, which minimizes a certain distance between the observed experimental data and the model observation that is computed with the BVP and that depends on the model parameter. In many cases, the analysis of such an inverse problem can have a unique solution in the framework of statistics, that is to say when the model parameter is modeled by a random quantity, with or without external noise on the model observation. In such a case, the random model observation is completely defined by its probability distribution (in finite or in infinite dimension) that is the unique transformation of the probability distribution of the random model parameter. This transformation is defined by the functional that maps the model parameter to the model observation. Such a formulation is constructed for obtaining a well-posed problem that has a unique solution in the probability theory framework. We refer the reader to [1] and [2] for an overview concerning the general methodologies for statistical and computational inverse problems, including general least-square inversion and the maximum likelihood method [3], and including the Bayesian approach [4, 5, 6, 7, 8, 9].

A non-Gaussian second-order random field is completely defined by its system of marginal probability distributions, which is an uncountable family of probability distributions on sets of finite dimension, and not only by its mean function and its covariance function as for a Gaussian random field. The experimental identification of such a non-Gaussian random field then requires the introduction of an adapted representation in order to be in capability to solve the statistical inverse problem. For any non-Gaussian second-order random field, an important type of representation is based on the use of the polynomial chaos expansion, for which the development and the use in computational sciences and engineering have been pioneered by Roger Ghanem in 1990-1991 [10]. An efficient construction has been proposed in [11, 12], which consists in combining a Karhunen-Loève expansion (that allows for doing a statistical reduced model) with a polynomial chaos expansion of the statistical reduced model. This type of construction has then been re-analyzed and used for solving boundary value problems using the spectral approach (see for instance [13, 14, 15, 16, 17, 18, 19, 20, 21]). The polynomial chaos expansion has also been extended for an arbitrary probability measure [22, 23, 24, 25, 26]. New algorithms have been proposed for obtaining a robust computation of realizations of high degrees polynomial chaos [27, 28]. The polynomial chaos expansion has also been extended to the case of the random coefficients [29], to the construction of a basis adaptation in homogeneous chaos spaces [30], and for a multimodal random vector [31].

The statistical inverse problem for identifying a non-Gaussian random field as a model parameter of a BVP, using polynomial chaos expansion has been initialized in [32, 33], used in [34, 35], and revisited in [36]. In [37], the construction of the probability model of the random

coefficients of the polynomial chaos expansion is proposed by using the asymptotic sampling Gaussian distribution constructed with the Fisher information matrix, and used for model validation [38]. This work have been developed for statistical inverse problems that are rather in low stochastic dimension, and new ingredients have been introduced in [39, 40, 28] for statistical inverse problems in high stochastic dimension. In using the reduced chaos decomposition with random coefficients of random fields [29], a Bayesian approach for identifying the posterior probability model of the random coefficients of the polynomial chaos expansion of the model parameter of the BVP has been proposed in [41] for the low stochastic dimension and in [42] for the high stochastic dimension. The experimental identification of a non-Gaussian positive matrix-valued random field in high stochastic dimension, using partial and limited experimental data for a model observation related to the random solution of a stochastic BVP, is a difficult problem that requires both adapted representations and methodologies [39, 40, 42, 43]. A complete development concerning advanced representations of non-Gaussian matrix-valued random fields and their identification by solving a statistical inverse problem can be found in [40, 43, 44, 45].

The present paper deals with the challenging problem related to the statistical inverse problem for the identification of a non-Gaussian tensor-valued random field in high stochastic dimension. Such a random field can, for instance, be the parameter of a boundary value problem. The available experimental data can be, either a large experimental data base for the random field itself that is to be identified, or can be a partial and limited experimental data base corresponding to an observation of the stochastic BVP. For the statistical inverse identification of non-Gaussian tensor-valued random fields, the methodology and the algorithms presented hereinafter are extracted from [39, 40, 28, 46, 47, 43, 44]. An important step in the methodology is the construction and the use of algebraic prior stochastic models (APSM) of such a non-Gaussian random field for which some advanced generators have been developed. The APSM and the associated generators presented and used hereinafter are those that have been published in [44, 45, 48, 49, 50, 51, 52, 53, 54, 55, 40, 56, 57, 58]. Three illustrations are presented: the stochastic modeling of track irregularities for high-speed trains and its experimental identification [59], the stochastic continuum modeling of random interphases from atomistic simulations for a polymer nanocomposite [60], and the multiscale identification of the random elasticity field at mesoscale of a heterogeneous microstructure using multiscale experimental observations [61, 62, 63].

Notations. The following notations are used:

$\mathbf{x} = (x_1, \ldots, x_d)$ is a vector in $\mathbb{R}^d$ with $d \geq 1$ an integer.
$\mathbb{M}_{n,m}(\mathbb{R})$ is the set of all the $(n \times m)$ real matrices.
$\mathbb{M}_n(\mathbb{R})$ is the set of all the $(n \times n)$ real matrices.
$\mathbb{M}_n^S(\mathbb{R})$ is the set of all the symmetric $(n \times n)$ real matrices.
$\mathbb{M}_n^+(\mathbb{R})$ is the set of all the positive-definite symmetric $(n \times n)$ real matrices.
$\mathbb{M}_n^+(\mathbb{R}) \subset \mathbb{M}_n^S(\mathbb{R}) \subset \mathbb{M}_n(\mathbb{R})$.
$[I_n]$ is the $(n \times n)$ real unity matrix.
tr$\{[A]\}$ is the trace of the square matrix $[A]$.
APSM: algebraic prior stochastic model.
BVP: boundary value problem.
KLE: Karhunen-Loëve expansion.
OAPSM: optimal algebraic prior stochastic model.
PCE: polynomial chaos expansion.

2 PROBLEM TO BE SOLVED AND DIFFICULTIES

2.1 Statistical inverse problem to be solved in high dimension

The statistical inverse problem consists in identifying a non-Gaussian $\mathbb{M}_n^+(\mathbb{R})$-valued random field $\{[\mathbb{A}(\mathbf{x})], \mathbf{x} \in \Omega\}$ indexed by a bounded subset Ω of $\mathbb{R}^d$, with d an integer greater or equal to 1. This identification is performed in using an experimental data set constituted of ν_{exp} deterministic vectors $\mathbf{u}^{\text{exp},1}, \ldots, \mathbf{u}^{\text{exp},\nu_{\text{exp}}}$ belonging to $\mathbb{R}^{n_u}$, which correspond to independent realizations of a random model observation $\mathbf{U}$ with values in $\mathbb{R}^{n_u}$ of a computational model, which is written as

$$\mathbf{U} = \mathbf{h}([\mathbb{A}(\mathbf{x}^1)], \ldots, [\mathbb{A}(\mathbf{x}^{N_p})]) \,, \tag{1}$$

in which $\mathbf{x}^1, \ldots, \mathbf{x}^{N_p}$ are N_p given points in Ω, and where $\mathbf{h}$ is a given deterministic nonlinear mapping from $\mathbb{M}_n^+(\mathbb{R}) \times \ldots \times \mathbb{M}_n^+(\mathbb{R})$ into $\mathbb{R}^{n_u}$. In general, mapping $\mathbf{h}$ is not explicitly known but is numerically constructed with the computational model. It should be noted that mapping $\mathbf{h}$ could be re-parameterized in replacing the set of random matrices $\{[\mathbb{A}(\mathbf{x}^k)], k = 1, \ldots, N_p\}$ by a random vector $\mathbf{Y}$ with values in an adapted subset $\mathcal{Y}$ of $\mathbb{R}^{n_s}$ with $n_s = N_p\, n(n+1)/2$.

2.2 An example of statistical inverse problem in high dimension

The experimental identification of the stochastic model of the elasticity field at mesoscale of a material for which the elastic heterogeneous microstructure cannot be described in terms of constituents, such as the cortical bone (see Figure 1) is an example of statistical inverse problem in high dimension. In such a case, the experimental identification of the non-Gaussian matrix-valued random field $\{[\mathbb{A}(\mathbf{x})], \mathbf{x} \in \Omega\}$ that models the apparent elasticity field at mesoscale is an interesting illustration.

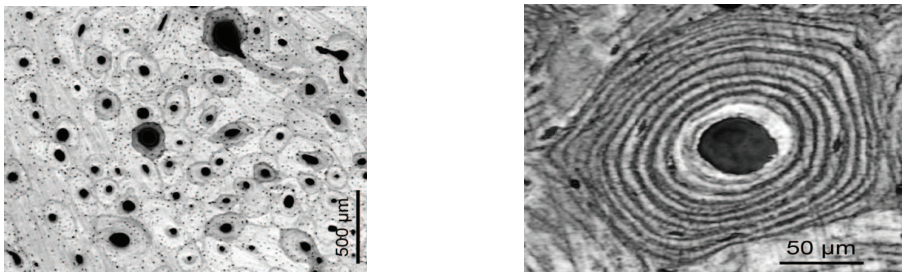

Figure 1: Cortical bone at mesoscale (left) and at microscale (right). Photo : Julius Wolff Institute, Charit - Universittsmedizin Berlin.

2.3 Difficulties of the statistical inverse problem in high dimension

The non-Gaussian second-order matrix-valued random field $\{\mathbb{A}(\mathbf{x}), \mathbf{x} \in \Omega\}$ must verify some important algebraic properties such as a deterministic or random boundedness, symmetry, positiveness, some invariance properties (for instance, induced by material symmetries), etc. In order to preserve a generality with respect to the chosen class in which the random field has to be identified, the polynomial chaos expansion (PCE) of the second-order random field is proposed, yielding a statistical inverse problem in high dimension. This is a challenging problem due to the high dimension and due to the fact that the random field is with values in a subset of $\mathbb{M}_n^+(\mathbb{R})$, which can be very complicated to describe and to explore for computing the PCE

coefficients from experimental data, and consequently, a direct approach cannot be used, and some adapted parameterizations and efficient algorithms must be developed.

3 METHODOLOGY FOR THE STATISTICAL INVERSE PROBLEM

For the statistical inverse problem related to the identification of a non-Gaussian vector- or matrix-valued random field in high dimension using partial and limited experimental data, the methodology consists in

(1) constructing an algebraic prior stochastic model of the random field and in identifying its hyperparameters using experimental data (Step 1);

(2) choosing an adapted representation for performing its polynomial chaos expansion in high dimension (Step 2);

(3) identifying a posterior stochastic model (Step 3).

This methodology is developed in details in [44, 45] for which the first developments and algorithms have been published in [39, 42, 40], for which additional complements concerning the algorithms that can be found in [27, 28], and for which mathematical developments concerning representation of random fields and mathematical analysis of stochastic elliptic boundary value problem are given in [43].

3.1 Step 1: Construction of an algebraic prior stochastic model (APSM) and its identification using experimental data

(i) The first stage in Step 1 concerns the construction of an adapted family, $\{[\mathbb{A}^{\text{APSM}}(\mathbf{x};\mathbf{b})], \mathbf{x} \in \Omega\}$, of algebraic prior stochastic models for the non-Gaussian $\mathbb{M}_n^+(\mathbb{R})$-valued random field $\{[\mathbb{A}(\mathbf{x})], \mathbf{x} \in \Omega\}$, defined on a probability space $(\Theta, \mathcal{T}, \mathcal{P})$, indexed by a bounded domain Ω in $\mathbb{R}^d$. This family of APSM depends on a vector-valued hyperparameter $\mathbf{b}$ belonging to an admissible set $\mathcal{B}_{\text{ad}}$ that is assumed to be in low dimension. For instance, such a hyperparameter is made up of the statistical mean matrix, some dispersion parameters, the spatial-correlation lengths, etc. An important example of construction of a APSM will be given after. It should be noted that in high dimension, the real possibility to correctly identify random field $[\mathbb{A}]$, through a stochastic boundary value problem, is directly related to the capability of the constructed APSM to represent fundamental properties such as lower bound, positiveness, invariance related to material symmetry, mean value, support of the spectrum, spatial-correlation lengths, level of statistical fluctuations, etc.

(ii) The second stage is related to the calculation of an optimal value, $\mathbf{b}^{\text{opt}}$, of hyperparameter $\mathbf{b}$ of the APSM using the maximum likelihood method for random model observation $\mathbf{U}$ of the computational model, and the experimental data set, $\mathbf{u}^{\text{exp},1}, \ldots, \mathbf{u}^{\text{exp},\nu_{\text{exp}}}$ for $\mathbf{U}$. The optimal value is written as

$$\mathbf{b}^{\text{opt}} = \arg \max_{\mathbf{b} \in \mathcal{B}_{\text{ad}}} \sum_{\ell=1}^{\nu_{\text{exp}}} \log p_{\mathbf{U}}(\mathbf{u}^{\text{exp},\ell}; \mathbf{b}) \,, \tag{2}$$

in which $\mathbf{u} \mapsto p_{\mathbf{U}}(\mathbf{u};\mathbf{b})$ is the probability density function of random model observation $\mathbf{U}$, estimated with the computational model $\mathbf{U} = \mathbf{h}([\mathbb{A}^{\text{APSM}}(\mathbf{x}^1;\mathbf{b})], \ldots, [\mathbb{A}^{\text{APSM}}(\mathbf{x}^{N_p};\mathbf{b})])$. The optimal PASM (OAPSM) of the random field $[\mathbb{A}]$ is denoted by $[\mathbb{A}^{\text{OAPSM}}]$ and is such that, for all $\mathbf{x}$ in Ω, $[\mathbb{A}^{\text{OAPSM}}(\mathbf{x})] =: [\mathbb{A}^{\text{APSM}}(\mathbf{x}, \mathbf{b}^{\text{opt}})]$. Finally, ν independent realizations of the OPASM can easily be computed using the APSM with $\mathbf{b} = \mathbf{b}^{\text{opt}}$,

$$\{[\mathbb{A}^{\text{OAPSM}}(\mathbf{x};\theta_\ell)], \mathbf{x} \in \Omega\} \quad , \quad \ell = 1, \ldots, \nu \,, \tag{3}$$

in which $\theta_1, \ldots, \theta_\nu$ are in Θ.

3.2 Step 2: Choosing an adapted representation for performing the polynomial chaos expansion in high dimension and identification of a prior stochastic model using experimental data

Step 2 is devoted to the construction of a polynomial chaos expansion of random field $[\mathbb{A}^{\text{OAPSM}}]$ that is defined by an optimal algebraic prior stochastic model. Such a PCE will be used for extending the PCE representation in the set of all the $\mathbb{M}_n^+(\mathbb{R})$-valued non-Gaussian second-order random fields by randomizing the coefficients (see Step3). However, such a PCE cannot be carried out with a direct approach because all the algebraic properties, such as the boundedness and and the positiveness of the random field for which the PCE should be performed, must be satisfied. Consequently, an indirect approach is performed consisting in constructing, in a first stage, an adapted representation for $\{[\mathbb{A}^{\text{OAPSM}}(\mathbf{x})], \mathbf{x} \in \Omega\}$ for which ν independent realizations have been computed (see Eq. (3)).

(i) The change of representation for random field $[\mathbb{A}^{\text{OAPSM}}]$ consists in introducing a class of lower-bounded random fields, normalized, and releasing some constraints, such that

$$[\mathbb{A}^{\text{OAPSM}}(\mathbf{x})] = \mathcal{G}([\mathbf{G}^{\text{OAPSM}}(\mathbf{x})]) \quad , \quad \forall\, \mathbf{x} \in \Omega\,, \tag{4}$$

in which $\mathcal{G}$ is an invertible transformation from $\mathbb{M}_n^S(\mathbb{R})$ into $\mathbb{M}_n^+(\mathbb{R})$. Consequently, from Eq. (3), ν independent realizations of the $\mathbb{M}_n^S(\mathbb{R})$-valued random field, $\{[\mathbf{G}^{\text{OAPSM}}(\mathbf{x})], \mathbf{x} \in \Omega\}$, can be deduced such that, for all $\mathbf{x}$ in Ω,

$$[\mathbf{G}^{\text{OAPSM}}(\mathbf{x}, \theta_\ell)] = \mathcal{G}^{-1}([\mathbb{A}^{\text{OAPSM}}(\mathbf{x}, \theta_\ell)]) \quad , \quad \ell = 1, \ldots, \nu\,. \tag{5}$$

Example of an invertible transformation. Some general invertible transformations and their mathematical analyses can be found in [40, 43, 44]. For instance, we can consider the following class of representation defined by

$$[\mathbb{A}^{\text{OAPSM}}(\mathbf{x})] = \frac{1}{1+\varepsilon}[\underline{L}(\mathbf{x})]^T \left\{\varepsilon[I_n] + \exp_{\mathbb{M}}([\mathbf{G}^{\text{OAPSM}}(\mathbf{x})])\right\}[\underline{L}(\mathbf{x})] \quad , \quad \forall \mathbf{x} \in \Omega\,, \tag{6}$$

in which ε is a positive real number, where $[\underline{L}(\mathbf{x})]$ is a deterministic $(n \times n)$ real matrix that is introduced for the normalization, and where $\exp_{\mathbb{M}}$ denotes the exponential for symmetric matrices, for which the inverse is the logarithm $\log_{\mathbb{M}}$ of positive-definite matrices. In such a class, the lower bound is the deterministic positive-definite matrix $[C_\ell(\mathbf{x})] = \frac{\varepsilon}{1+\varepsilon}[\underline{L}(\mathbf{x})]^T[\underline{L}(\mathbf{x})]$, and for all random matrix $[\mathbf{G}^{\text{OAPSM}}(\mathbf{x})]$ with values in $\mathbb{M}_n^S(\mathbb{R})$, the random matrix $[\mathbb{A}^{\text{OAPSM}}(\mathbf{x})] - [C_\ell(\mathbf{x})]$ is almost surely positive definite.

(ii) It is now possible to perform a PCE of the $\mathbb{M}_n^S(\mathbb{R})$-valued random field $\{[\mathbf{G}^{\text{OAPSM}}(\mathbf{x})], \mathbf{x} \in \Omega\}$. The general methodology consists in performing a Karhunen-Loëve expansion (KLE) of the random field, followed by a PCE of the coordinates of the KLE (see [10, 12]). Concerning a new approach devoted to the KLE of vector-valued random fields with a scaling and the use of an optimal basis, we refer the reader to [46, 47], and for the KLE of matrix-valued random fields, to [40, 43, 44]. It should be noted that, in the framework of statistical inverse problem with partial and limited experimental data, the covariance operator of random field $[\mathbb{A}]$ cannot be estimated with nonparametric statistics due to the lack of data and consequently, the KLE of the random

field cannot be constructed by this way. However, with the methodology that is proposed, the covariance operator of random field $[\mathbf{G}^{\text{OAPSM}}]$ can be estimated using the realizations computed with Eq. (5). The truncated KLE of random field $[\mathbf{G}^{\text{OAPSM}}]$ is constructed at the order m, and then the coordinates of this expansion are expressed as a truncated PCE. Assuming that $[\mathbf{G}^{\text{OAPSM}}]$ is a second-order and mean-square continuous random field, then the truncated KLE at order m is written as,

$$[\mathbf{G}^{\text{OAPSM}\,(m)}(\mathbf{x})] = [G_0(\mathbf{x})] + \sum_{i=1}^{m} \sqrt{\lambda_i}\,[G_i(\mathbf{x})]\,\eta_i^{\text{OAPSM}} \quad , \quad \forall \mathbf{x} \in \Omega\,, \tag{7}$$

in which, for all $\mathbf{x}$ in Ω, the random matrix $[G_0(\mathbf{x})] = E\{[\mathbf{G}^{\text{OAPSM}}(\mathbf{x})]\}$ is the mean value in $\mathbb{M}_n^S(\mathbb{R})$, and where $\boldsymbol{\eta}^{\text{OAPSM}} = (\eta_1^{\text{OAPSM}}, \ldots, \eta_m^{\text{OAPSM}})$ is a non-Gaussian random vector with values in $\mathbb{R}^m$, which is centered and for which its covariance matrix is $[C_{\boldsymbol{\eta}^{\text{OAPSM}}}] = [I_m]$ (the components of $\boldsymbol{\eta}^{\text{OAPSM}}$ are uncorrelated and normalized, but are statistically dependent). From Eqs. (5) and (7), ν independent realizations of random vector $\boldsymbol{\eta}^{\text{OAPSM}}$ are computed,

$$\eta_i^{\text{OAPSM}}(\theta_\ell) = \frac{1}{\sqrt{\lambda_i}} \int_\Omega \text{tr}\{[G_i(\mathbf{x})]^T\,([\mathbf{G}^{\text{OAPSM}}(\mathbf{x},\theta_\ell)] - [G_0(\mathbf{x})])\}\,d\mathbf{x} \quad , \quad \ell = 1, \ldots, \nu\,. \tag{8}$$

The truncated PCE of the non-Gaussian random vector $\boldsymbol{\eta}^{\text{OAPSM}}$ is denoted by $\boldsymbol{\eta}^{\text{chaos}}(N, N_g)$ and is written as,

$$\boldsymbol{\eta}^{\text{OAPSM}} \simeq \boldsymbol{\eta}^{\text{chaos}}(N, N_g) = [z]^T\,\boldsymbol{\Psi}(\boldsymbol{\Xi})\,, \tag{9}$$

in which the matrix $[z]$ of the coefficients belongs to the compact Stiefel manifold $\mathbb{V}_m(\mathbb{R}^N) \subset \mathbb{M}_{N,m}$ that is defined by

$$\mathbb{V}_m(\mathbb{R}^N) = \left\{[z] \in \mathbb{M}_{N,m}(\mathbb{R})\,;\; [z]^T\,[z] = [I_m]\right\}\,. \tag{10}$$

In Eq. (9), the random vector $\boldsymbol{\Xi} = (\Xi_1, \ldots, \Xi_{N_g})$ with values in $\mathbb{R}^{N_g}$ has a probability distribution $P_{\boldsymbol{\Xi}}(d\boldsymbol{\xi}) = p_{\boldsymbol{\Xi}}(\boldsymbol{\xi})\,d\boldsymbol{\xi}$ (on $\mathbb{R}^{N_g}$) that is assumed to be given and that must verify,

$$\int_{\mathbb{R}^{N_g}} \|\boldsymbol{\xi}\|^m\,p_{\boldsymbol{\Xi}}(\boldsymbol{\xi})\,d\boldsymbol{\xi} < +\infty \quad , \quad \forall\, m \in \mathbb{N}\,.$$

The random vector $\boldsymbol{\Psi}(\boldsymbol{\Xi}) = (\Psi_{\boldsymbol{\alpha}^{(1)}}(\boldsymbol{\Xi}), \ldots, \Psi_{\boldsymbol{\alpha}^{(N)}}(\boldsymbol{\Xi}))$ is made up of the multivariate polynomials (the chaos) that are orthonormal with respect to $P_{\boldsymbol{\Xi}}(d\boldsymbol{\xi})$,

$$E\{\Psi_{\boldsymbol{\alpha}^{(j)}}(\boldsymbol{\Xi})\,\Psi_{\boldsymbol{\alpha}^{(k)}}(\boldsymbol{\Xi})\} = \int_{\mathbb{R}^{N_g}} \Psi_{\boldsymbol{\alpha}^{(j)}}(\boldsymbol{\xi})\,\Psi_{\boldsymbol{\alpha}^{(k)}}(\boldsymbol{\xi})\,p_{\boldsymbol{\Xi}}(\boldsymbol{\xi})\,d\boldsymbol{\xi} = \delta_{jk}\,. \tag{11}$$

If $N_d \geq 1$ is the maximum degree of the polynomials, then the integer N is given by

$$N = \frac{(N_g + N_d)!}{N_g!\,N_d!} - 1\,. \tag{12}$$

$\diamond$ If $\boldsymbol{\Xi}$ is the normalized Gaussian random vector, then $\Psi_{\boldsymbol{\alpha}^{(j)}}(\boldsymbol{\Xi}) = (\Phi_{\alpha_1^{(j)}}(\Xi_1) \times \ldots \times \Phi_{\alpha_{N_g}^{(j)}}(\Xi_{N_g}))$ in which $\Phi_{\alpha_k^{(j)}}(\Xi_k)$ is the usual normalized Hermite polynomial on $\mathbb{R}$ (see [10, 12]).

$\diamond$ If the Gaussian measure is replaced by an arbitrary probability measure for which the real-valued random variables $\Xi_1, \ldots, \Xi_{N_g}$ are statistically independent, see [22].

◊ For an arbitrary measure for which $\Xi_1, \ldots, \Xi_{N_g}$ are statistically dependent, see [31].

For these three cases, if the degree N_d of the polynomials is high, then the realizations $\boldsymbol{\Psi}(\boldsymbol{\Xi}(\theta_\ell)$ cannot be computed with usual algorithms. An efficient algorithm must be used for obtaining a robust computation of realizations for the high degrees polynomial chaos. For the details of such an algorithm, see [27, 28, 31].

(iii) The next stage consists in calculating (1) the coefficients that are represented by matrix $[z]$ and (2) the optimal values of N_g and N in order to obtain an error that is sufficiently small between $\boldsymbol{\eta}^{\text{OAPSM}}$ and $\boldsymbol{\eta}^{\text{chaos}}(N, N_g)$. For each fixed value of N and N_g, an optimal value $[z_0(N, N_g)] \in \mathbb{V}_m(\mathbb{R}^N) \subset \mathbb{M}_{N,m}$ of $[z]$ is computed with the maximum likelihood method,

$$[z_0(N, N_g)] = \arg \max_{[z]\in\mathbb{V}_m(\mathbb{R}^N)} \sum_{\ell=1}^{\nu} \log p_{\boldsymbol{\eta}^{\text{chaos}}(N,N_g)}(\boldsymbol{\eta}^{\text{OAPSM}}(\theta_\ell)\,;\,[z])\,, \tag{13}$$

in which the realizations $\boldsymbol{\eta}^{\text{OAPSM}}(\theta_\ell)$ are given by Eq. (8) and where the probability density function $\mathbf{y} \mapsto p_{\boldsymbol{\eta}^{\text{chaos}}(N,N_g)}(\mathbf{y}\,;\,[z])$ of random vector $\boldsymbol{\eta}^{\text{chaos}}(N, N_g)$ is estimated at point $\mathbf{y} = \boldsymbol{\eta}^{\text{OAPSM}}(\theta_\ell)$ by the multidimensional kernel density estimator and by using a great number of independent realizations that are generated with the stochastic model defined by Eq. (9). The optimization problem defined by Eq. (13) is solved with the efficient random search algorithm on the manifold $\mathbb{V}_m(\mathbb{R}^N)$, which is detailed in [39, 28]. The optimal values N^{opt} and N_g^{opt} of N and N_g can be computed as explained in [28], which corresponds to an improvement of the following one that consists in minimizing the error function $(N, N_g) \mapsto \text{err}(N, N_g)$ defined by

$$\text{err}(N, N_g) = \frac{1}{m} \sum_{i=1}^{m} \int_{\text{BI}_i} |\log_{10} p_{\eta_i^{\text{OAPSM}}}(e) - \log_{10} p_{\eta_i^{\text{chaos}}(N,N_g)}(e\,;\,[z_0(N, N_g)])|\, de\,, \tag{14}$$

in which BI_i is the support of the kernel density estimator of $p_{\eta_i^{\text{OAPSM}}}$. The error function defined by Eq. (14) allows the accuracy of the tail of the distributions to be quantified. In order to simplify the notation, the optimal values N^{opt} and N_g^{opt} are re-written as N and N_g. Once the optimal value $[z_0(N, N_g)]$ has been computed, the optimal representation $\{[\mathbb{A}^{\text{OAPSM}}(\mathbf{x})], \mathbf{x} \in \Omega\}$ of the random field $\{[\mathbb{A}(\mathbf{x})], \mathbf{x} \in \Omega\}$ can then be deduced:

$$[\mathbb{A}^{\text{OAPSM}}(\mathbf{x})] = \mathcal{A}^{(m,N,N_g)}(\mathbf{x}, \boldsymbol{\Xi}, [z_0]) \quad , \quad \forall\, \mathbf{x} \in \Omega\,, \tag{15}$$

in which the mapping $(\mathbf{x}, \boldsymbol{\Xi}, [z_0]) \mapsto \mathcal{A}^{(m,N,N_g)}(\mathbf{x}, \boldsymbol{\Xi}, [z_0])$ is explicitly defined by Eqs. (4), (7), and (9), in which $[z_0]$ means $[z_0(N, N_g)]$.

(iv) The last stage of Step 2 consists in performing the experimental identification of the prior stochastic model $\{[\mathbb{A}^{\text{prior}}(\mathbf{x})], \mathbf{x} \in \Omega\}$ of random field $\{[\mathbb{A}(\mathbf{x})], \mathbf{x} \in \Omega\}$, which is defined by

$$[\mathbb{A}^{\text{prior}}(\mathbf{x})] = \mathcal{A}^{(m,N,N_g)}(\mathbf{x}, \boldsymbol{\Xi}, [z^{\text{prior}}])\,, \tag{16}$$

in which $[z^{\text{prior}}] \in \mathbb{V}_m(\mathbb{R}^N) \subset \mathbb{M}_{N,m}$ is computed with the maximum likelihood method for random model observation $\mathbf{U}$ of the computational model, using the experimental data set $\mathbf{u}^{\text{exp},1}, \ldots, \mathbf{u}^{\text{exp},\nu_{\text{exp}}}$ relative to $\mathbf{U}$. We then have

$$[z^{\text{prior}}] = \arg \max_{[z]\in\mathbb{V}_m(\mathbb{R}^N)} \sum_{\ell=1}^{\nu_{\text{exp}}} \log p_{\mathbf{U}^{(m,N,N_g)}}(\mathbf{u}^{\text{exp},\ell}; [z])\,, \tag{17}$$

in which $\mathbf{u} \mapsto p_{\mathbf{U}^{(m,N,N_g)}}(\mathbf{u}; [z])$ is the probability density function of random model observation $\mathbf{U}$, estimated with the computational model $\mathbf{U} = \mathbf{h}([\mathbb{A}^{\text{prior}}(\mathbf{x}^1)], \ldots, [\mathbb{A}^{\text{prior}}(\mathbf{x}^{N_p})])$ in which $[\mathbb{A}^{\text{prior}}(\mathbf{x}^k)] = \mathcal{A}^{(m,N,N_g)}(\mathbf{x}^k, \boldsymbol{\Xi}, [z^{\text{prior}}])$ with $k = 1, \ldots, N_p$. The optimization problem defined by Eq. (17) is solved with a random search algorithm that explores the neighborhood of a subset of $\mathbb{V}_m(\mathbb{R}^N)$ centered in $[z_0]$. For that, a parameterization

$$[z] = \mathcal{R}_{[z_0]}([w], \sigma) \tag{18}$$

of $\mathbb{V}_m(\mathbb{R}^N)$ (based on [64]) depending of a positive parameter σ, is used. The function $[w] \mapsto [z] = \mathcal{R}_{[z_0]}([w], \sigma)$ is a mapping from $\mathbb{M}_{N,m}(\mathbb{R})$ onto $\mathbb{V}_m(\mathbb{R}^N)$ such that $\mathcal{R}_{[z_0]}([0], \sigma) = [z_0]$ and, if $[w]$ belongs to a subset of $\mathbb{M}_{N,m}(\mathbb{R})$ centered in $[w] = [0]$ and with a sufficiently small diameter controlled by σ, then $[z] = \mathcal{R}_{[z_0]}([w], \sigma)$ belongs to a subset of $\mathbb{V}_m(\mathbb{R}^N)$ approximatively centered in $[z] = [z_0]$. This parameterization allows for exploring a subset of $\mathbb{V}_m(\mathbb{R}^N)$, centered in $[z_0] \in \mathbb{V}_m(\mathbb{R}^N)$, and for which its "diameter" is controlled by σ.

3.3 Step 3: Identifying a posterior stochastic model

For constructing a posterior stochastic model $\{[\mathbb{A}^{\text{post}}(\mathbf{x})], \mathbf{x} \in \Omega\}$ of random field $\{[\mathbb{A}(\mathbf{x})], \mathbf{x} \in \Omega\}$, matrix $[z]$ of the PCE is modeled by a random matrix $[\mathbf{Z}]$ with values in $\mathbb{V}_m(\mathbb{R}^N)$, for which the statistical fluctuations are in a subset of $\mathbb{V}_m(\mathbb{R}^N)$ centered around $[z^{\text{prior}}]$ computed in Step 2. A prior stochastic model $[\mathbf{Z}^{\text{prior}}]$ of $[\mathbf{Z}]$ is thus introduced in using the representation defined by Eq. (18),

$$[\mathbf{Z}^{\text{prior}}] = \mathcal{R}_{[z^{\text{prior}}]}([\mathbf{W}^{\text{prior}}], \sigma) \tag{19}$$

in which $[\mathbf{W}^{\text{prior}}]$ is a centered and normalized random matrix whose probability distribution and its generator are known. For a sufficiently small value of σ, the statistical fluctuations of the $\mathbb{V}_m(\mathbb{R}^N)$-valued random matrix $[\mathbf{Z}^{\text{prior}}]$ are approximatively centered around $[z^{\text{prior}}]$. The Bayesian update allows the posterior distribution of the random matrix $[\mathbf{W}^{\text{post}}]$ with values in $\mathbb{M}_{N,m}(\mathbb{R})$ to be estimated using the experimental data set $\mathbf{u}^{\text{exp},1}, \ldots, \mathbf{u}^{\text{exp},\nu_{\text{exp}}}$ and the prior stochastic model of the random model observation $\mathbf{U} = \mathbf{h}([\mathbb{A}^{\text{prior,prior}}(\mathbf{x}^1)], \ldots, [\mathbb{A}^{\text{prior,prior}}(\mathbf{x}^{N_p})])$ in which $[\mathbb{A}^{\text{prior,prior}}(\mathbf{x}^k)] = \mathcal{A}^{(m,N,N_g)}(\mathbf{x}^k, \boldsymbol{\Xi}, [\mathbf{Z}^{\text{prior}}])$ with $k = 1, \ldots, N_p$. The posterior stochastic model $[\mathbb{A}^{\text{post}}]$ of random field $[\mathbb{A}]$ is then given by

$$[\mathbb{A}^{\text{post}}(\mathbf{x})] = \mathcal{A}^{(m,N,N_g)}(\mathbf{x}, \boldsymbol{\Xi}, \mathcal{R}_{[z^{\text{prior}}]}([\mathbf{W}^{\text{post}}])) \quad , \quad \forall \mathbf{x} \in \Omega\,. \tag{20}$$

The identification procedure can then be restarted from Step 2 if necessary for improving the identification.

4 ALGEBRAIC PRIOR STOCHASTIC MODEL

In this section, we present some elements related to the development of an algebraic prior stochastic model for the non-Gaussian positive-definite matrix-valued random fields with statistical fluctuations in a given symmetry class and in the anisotropic class. All the details of this development can be found in [48, 49] concerning the stochastic model for the anisotropic statistical fluctuations. The case for which dominant statistical fluctuations belong to the isotropic class with anisotropic statistical fluctuations has been introduced in [65] and has been generalized to all the symmetry classes in [56] and included references. A very general theory is then proposed in [56, 58, 45, 44] for the general case for which there are simultaneously some statistical fluctuations in a given symmetry class and statistical fluctuations in the anisotropic class.

4.1 Algebraic prior stochastic models for non-Gaussian matrix-valued random field

The construction of an algebraic prior stochastic model for a non-Gaussian $\mathbb{M}_n^+(\mathbb{R})$-valued random field $\{[\mathbb{A}^{\text{APSM}}(\mathbf{x})], \mathbf{x} \in \Omega\}$ is defined hereinafter in the framework of the 3D linear elasticity ($n = 6$). A symmetry class, induced by a material symmetry, is defined as a subset $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$ of $\mathbb{M}_n^+(\mathbb{R})$, whose algebraic dimension is $n_s \leq n(n+1)/2$ (see its construction in [56]). For instance, $n_s = 2$ for the isotropy, $n_s = 5$ for the transverse isotropy, and $n_s = 21$ for the anisotropy. For all $\mathbf{x}$ fixed in Ω, the $\mathbb{M}_n^+(\mathbb{R})$-valued random matrix $[\mathbb{A}^{\text{APSM}}(\mathbf{x})]$ is, in mean, close to a given symmetry class $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$ that is independent of $\mathbf{x}$ and exhibits dominant statistical fluctuations in the symmetry class $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$ with some anisotropic statistical fluctuations in $\mathbb{M}_n^+(\mathbb{R})$ (around this symmetry class), which is controlled independently of the level of statistical fluctuations in the symmetry class.

4.2 Representation of the random field

The algebraic representation of the random field $\{[\mathbb{A}^{\text{APSM}}(\mathbf{x})], \mathbf{x} \in \Omega\}$ is constructed such that

$$[\mathbb{A}^{\text{APSM}}(\mathbf{x})] = [C_\ell] + [\mathbf{A}(\mathbf{x})] \quad , \quad \forall\, \mathbf{x} \in \Omega\,, \tag{21}$$

in which $[C_\ell]$ is the $\mathbb{M}_n^+(\mathbb{R})$-valued deterministic lower bound assuring the uniform ellipticity for the stochastic elliptic operator constructed with random field $[\mathbb{A}^{\text{APSM}}]$. The non-Gaussian $\mathbb{M}_n^+(\mathbb{R})$-valued random field $\{[\mathbf{A}(\mathbf{x})], \mathbf{x} \in \Omega\}$ is written as

$$[\mathbf{A}(\mathbf{x})] = [\underline{S}]^T\, [\mathbf{M}(\mathbf{x})]^{1/2} [\mathbf{G}_0(\mathbf{x})]\, [\mathbf{M}(\mathbf{x})]^{1/2}\, [\underline{S}]\,, \tag{22}$$

in which $\{[\mathbf{M}(\mathbf{x})], \mathbf{x} \in \Omega\}$ is a non-Gaussian $\mathbb{M}^{\text{sym}}(\mathbb{R})$-valued random field that models the dominant statistical fluctuations in the symmetry class defined by $\mathbb{M}^{\text{sym}}(\mathbb{R})$ (see its construction and its generator in [56]), where $\{[\mathbf{G}_0(\mathbf{x})], \mathbf{x} \in \Omega\}$ is a non-Gaussian $\mathbb{M}_n^+(\mathbb{R})$-valued random field that models the anisotropic statistical fluctuations around the symmetry class (see its construction and its generator in [48, 49]). The random fields $\{[\mathbf{M}(\mathbf{x})], \mathbf{x} \in \Omega\}$ and $\{[\mathbf{G}_0(\mathbf{x})], \mathbf{x} \in \Omega\}$ are statistically independent. The invertible deterministic matrix $[\underline{S}]$ that belongs to $\mathbb{M}_n(\mathbb{R})$ is introduced for the normalization (see its construction in [56, 44]). It should be noted that the generator of independent realizations of random field $\{[\mathbf{M}(\mathbf{x})], \mathbf{x} \in \Omega\}$ proposed in [56, 44] is based on an efficient MCMC algorithm constructed in using a nonlinear Itô stochastic differential equation associated with a second-order Hamiltonian dissipative nonlinear dynamical system.

5 APPLICATION 1

This application deals with the stochastic modeling in high dimension of track irregularities, for nonlinear stochastic dynamic simulations of high speed trains. All the details and the results given for this application are extracted form the work published in [28, 46, 47, 59, 66]. The objectives are the following. The first one consists in performing the experimental identification of a stochastic modeling of the track geometry for portions of same length, using a set of realizations corresponding to the experimental measurements of the track geometry for several TGV lines (high speed train lines). The second one is the validation of the identified stochastic model of the track geometry by using a comparison between the experimental measurements of the dynamical responses of a TGV train and the simulations performed with the nonlinear stochastic dynamical model of the TGV train for which the track geometry is the identified stochastic model.

5.1 Description of track irregularities

The irregularities of the track geometry are parameterized by 4 parameters, $X_1(s)$, $X_2(s)$, $X_3(s)$, and $X_4(s)$, as a function of the curvilinear abscissa s (see Figure 2-left). As an example, Figure 2-right displays the values of the irregularities for the 4 parameters. The track portions have a same length S with $S < 10,000\,m$ and are sampled with a spatial step Δs. Consequently, the track geometry is represented by a random vector $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4)$ with values in $N_S = 4 \times (S/\Delta s + 1)$ in which $\mathbf{X}_j = (X_j(0), X_j(\Delta s), \ldots, X_j(S))$. Random vector $\mathbf{X}$ is in high dimension (N_S greater than $10,000$). The experimental measurements of the track geometry have been carried out for $\nu^{\text{exp}} = 1,730$ track portions of length S with spatial sampling Δs yielding ν^{exp} independent realizations $\mathbf{x}^1, \ldots, \mathbf{x}^{\nu^{\text{exp}}}$ in $\mathbb{R}^{N_S}$ of $\mathbf{X}$.

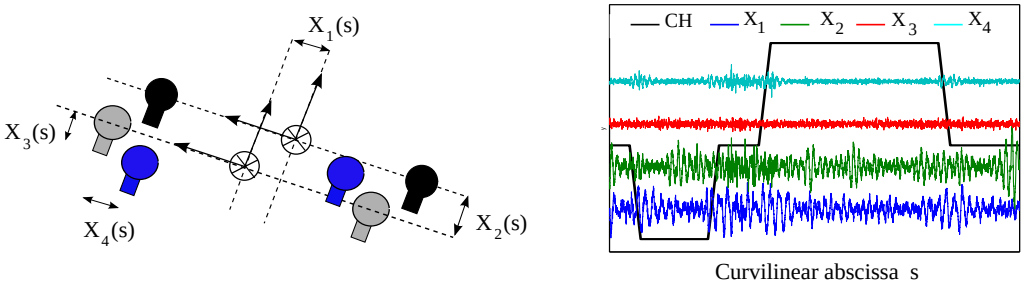

Figure 2: Parameterization of the irregularities of the track geometry (left figure). Example of track irregularities for the four parameters $X_1(s), X_2(s), X_3(s), X_4(s)$ as a function of curvilinear abscissa s and CH is the horizontal curvature.

5.2 Stochastic modeling and experimental identification of the track irregularities

The stochastic modeling of the track irregularities and its experimental identification are performed following the methodology described hereinbefore and are detailed in [28, 66]. The estimations $\underline{\widehat{\mathbf{X}}}$ of the mean value and $[\widehat{C}_{\mathbf{X}}]$ of the covariance matrix of random vector $\mathbf{X}$ are calculated using the independent realizations $\mathbf{x}^1, \ldots, \mathbf{x}^{\nu^{\text{exp}}}$ with $\nu^{\text{exp}} = 1,730$. The principal component analysis allows for writing the approximation $\mathbf{X} \simeq \underline{\widehat{\mathbf{X}}} + \sum_{i=1}^{m} \sqrt{\lambda_i}\, \eta_i\, \mathbf{w}^i$ with $m = 940$ that yields a relative error less than 1%. The polynomial chaos expansion of the random vector $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_m)$ with values in $\mathbb{R}^m$ is written as $\boldsymbol{\eta} \simeq \boldsymbol{\eta}^{\text{chaos}}(N, N_g) = \sum_{j=1}^{N} \mathbf{y}^j\, \Psi_{\boldsymbol{\alpha}^{(j)}}(\boldsymbol{\Xi})$ in which the components of the random vector $\boldsymbol{\Xi} = (\Xi_1, \ldots \Xi_{N_g})$ are independent random variables uniformly distributed on $[-1\,,1]$ and where the normalized multivariate polynomials are written as $\Psi_{\boldsymbol{\alpha}^{(j)}}(\boldsymbol{\Xi}) = \Phi_{\alpha_1^{(j)}}(\Xi_1) \times \ldots \times \Phi_{\alpha_{N_g}^{(j)}}(\Xi_{N_g})$ with $\Phi_{\alpha_k^{(j)}}(\Xi_k)$ the normalized univariate Legendre polynomial. The integers N and N_g are identified in the statistical inverse procedure as explained in Section 3. The vector-valued coefficients $\mathbf{y}^j$ for $j = 1, \ldots, N$ are estimated using the maximum likelihood method for which an algorithm adapted to the high dimension is used (see Section 3). The error is estimated with the following error function depending on N and N_g, which is written as $\text{err}(N, N_g) = \sum_{i=1}^{m} \int_{\text{BI}_i} |\log_{10} p_{\eta_i}(e) - \log_{10} p_{\eta_i^{\text{chaos}}(N,N_g)}(e\,;\, \mathbf{y}^1, \ldots, \mathbf{y}^N)|\, de$, in which BI_i is the interval bounding the experimental values η_i^{exp}. Figure 3 displays the graph of function $N \mapsto \text{err}(N, N_g)$ for $N_g = 3$ and $N_g = 4$. The convergence is reached for $N_g = 3$ and $N = 2,925$.

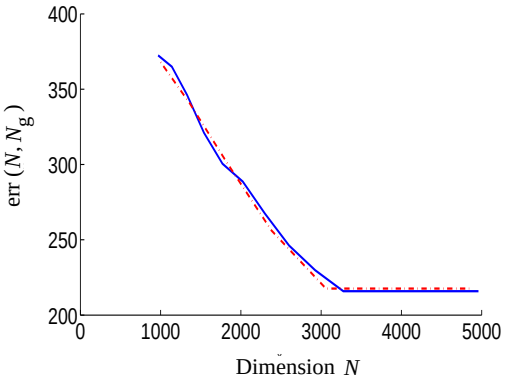

Figure 3: Graph of function $N \mapsto \text{err}(N, N_g)$ for $N_g = 3$ (solid line) and $N_g = 4$ (dash-dotted line).

5.3 Validation using the nonlinear dynamical responses of a TGV train

A validation of the identified stochastic model of the track geometry can be obtained in comparing the experimental measurements and the simulations for the dynamical responses of a TGV train. A track of $5\,km$ has been considered for which the track geometry has been measured and for which the dynamical responses of a TGV train has also been measured. The simulation of the nonlinear stochastic dynamics of the TGV train exited by the general stochastic model of the whole track geometry (which has been experimentally identified as described in the previous section). Figure 4 displays the comparisons between experimental measurements and the simulations. Two types of analysis are presented. The first one is related to a statistical analysis in comparing the mean number of upcrossings. The second one deals with a spectral analysis in comparing the power spectral density function. The observed responses of the train are 4 transverse contact forces between rail and wheel, C_1, C_2, C_3 and C_4, at wheelsets of bogies in the TGV train. The experimental measurements are denoted by C_j^{exp} and the corresponding simulated quantities are denoted by C_j^{sim}. It can be seen that the comparisons are really good for mean number of upcrossings and for the power spectral density function.

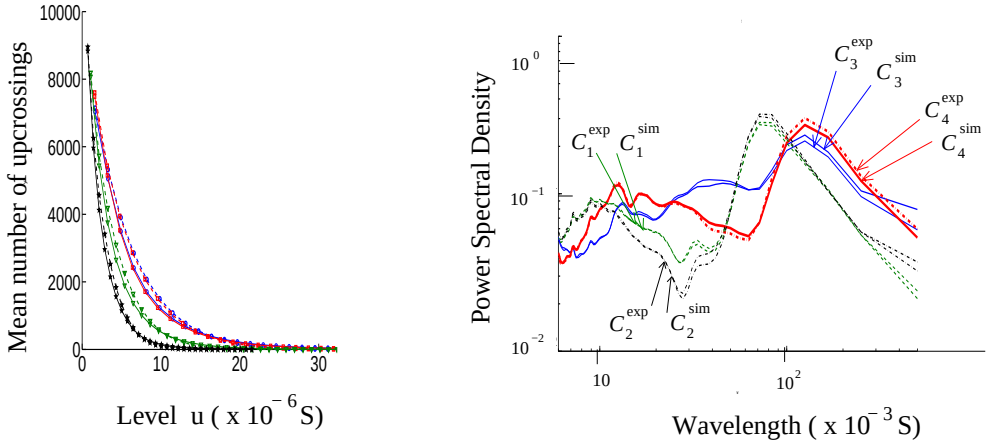

Figure 4: Comparisons between experimental measurements C_j^{exp} and simulation C_j^{sim} for 4 transverse contact forces. Mean number of upcrossings (left figure). Power spectral density function (right figure).

6 APPLICATION 2

This application is devoted to the stochastic continuum modeling of random interphases from atomistic simulations for a polymer nanocomposite. All the details and results given for describing this application are extracted form the work published in [60]. The objective is to construct, in the framework of continuum mechanics, a prior stochastic model of the matrix-valued random elasticity field $\mathbf{x} \mapsto [\mathbb{A}^{\text{meso}}(\mathbf{x})]$ describing the random elastic behavior of the interphase between an amorphous polymer and a silicon nano-inclusion inserted in the polymer, and to identify the hyperparameters of the prior stochastic model of $[\mathbb{A}^{\text{meso}}]$ by using the atomistic simulations that are considered as some simulated experiments by solving a statistical inverse problem.

6.1 Physical description, molecular dynamics modeling, and simulation procedure

The polymer nanocomposite is made up of an amorphous polymer containing one silicon nano-inclusion. The amorphous polymer is constituted of long chains with CH_2 sites, represented through a coarse graining with stiffness potentials and the Lennard Jones potential. The silicon nano-inclusion is an amorphous bulk of SiO_2 molecules for which a fully atomistic description is done in terms of Si and O atoms, for which the Coulomb potential is used. The interaction CH_2 - SiO_2 is taken into account by the Lennard Jones potential. As an illustration, Figure 5 displays 3D and 2D views.

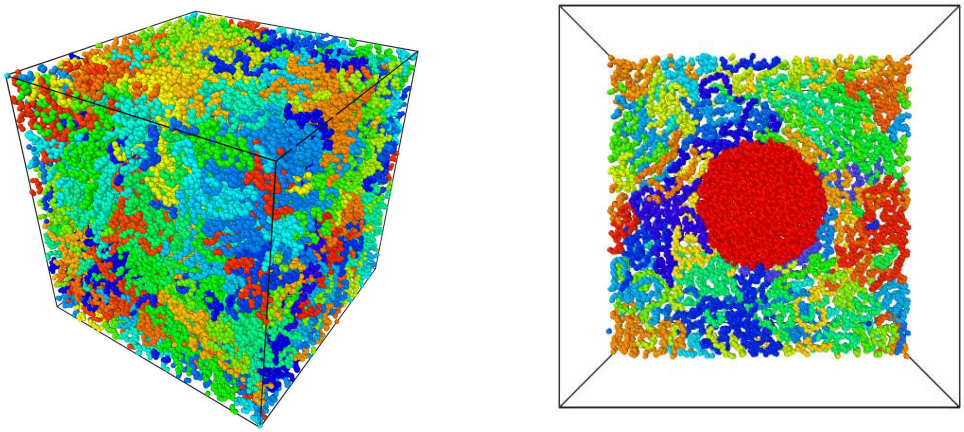

Figure 5: Example: 3D view of the distribution for 80 polymer chains (left figure) and a 2D view showing the SiO_2 nano-inclusion (right figure), for a domain that is a cube with a side of $13.5 \times 10^{-9}\, m$ in which the silicon sphere has a diameter $6 \times 10^{-9}\, m$.

All the simulations are carried out with a target volume fraction of 4.7%. For computation, the simulation domain is a cube with $6.8 \times 10^{-9}\, m$ side and contains 10 polymer chains. Each polymer chain has $1,000$ CH_2 sites yielding a total of $10,000$ CH_2 for the polymer bulk. The SiO_2 nano-inclusion is a sphere with $3 \times 10^{-9}\, m$ diameter. The number of Si atoms is 275 and the the number of O atoms is 644.

The simulation procedure is constituted of two main steps: the first one is related to atomistic simulations of the polymer nanocomposite and the second one to the statistical inverse problem for the identification of a prior continuum stochastic model of the random elasticity field that describes the random elastic behavior of the interphase.

6.2 Atomistic simulations with 10 chains

Atomistic simulations are done for a set of configurations in order to generate a set of realizations that are used for identifying the prior stochastic continuum model of the random elasticity field of the interphase (in the framework of continuum mechanics). For the simulations, the temperature is $T = 100\,^oK$ and the pressure P is the control variable. The adapted conditions are introduced for simulating 6 traction and shear mechanical tests. A time-spatial averaging is computed in order to estimate the apparent strain that allows for deducing the components of the apparent elasticity matrix in the sense of continuum mechanics. The simulations have been done with the molecular dynamics software, Lammps, for 20 configurations that have required 4 months of computation with 48 cores. The simulations have allowed for identifying the interphase thickness, e, which is about $2 \times 10^{-9}\, m$. An analysis with respect to the diameter of the silicon nano-inclusion (3, 6, and $9.6\, nm$) shows that the interphase thickness is independent of the diameter.

Example of a partial result obtained with the atomistic simulation. As an example, Figure 6 displays the polymer density ρ^n in the nano-composite divided by the pure polymer density ρ^p as a function of the distance r from the center of the sphere representing the silicon nano-inclusion.

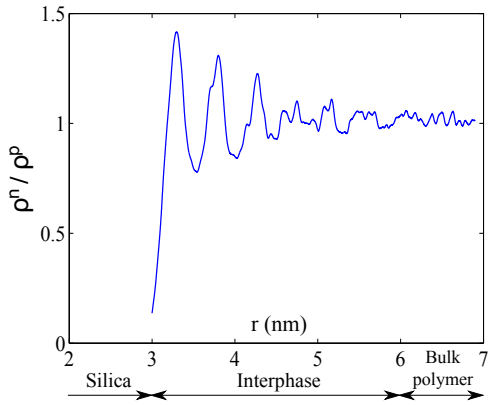

Figure 6: Polymer density ρ^n in the nano-composite divided by the pure polymer density ρ^p as a function of the distance r in nanometer from the center of the sphere representing the silicon nano-inclusion for the case of 80 polymer chains in a domain that is a cube with a side of $13.5 \times 10^{-9}\, m$ in which the silicon nano-inclusion sphere has a diameter $6 \times 10^{-9}\, m$.

6.3 Statistical inverse problem for the identification of a prior stochastic continuum model of the interphase

The prior stochastic model of the non-Gaussian matrix-valued random elasticity field $\mathbf{x} \mapsto [\mathbb{A}^{\text{meso}}(\mathbf{x})]$ is chosen is the class of transversally isotropic material symmetries [56]. Using the spherical coordinates, there are 5 dependent random fields that have to be identified as previously explained. The hyperparameters that have to be identified are the dispersion parameter allowing for controlling the statistical fluctuations level and the spatial-correlation lengths. The finite element method is used for solving the 6 stochastic boundary value problems corresponding to the 6 mechanical tests. For the system with 10 polymer chains, there are $190,310$ four nodes finite elements with one Gauss point, $34,187$ nodes and $102,561$ degrees of freedom

(see Figure 7). The maximum likelihood method is used for estimating the optimal values of

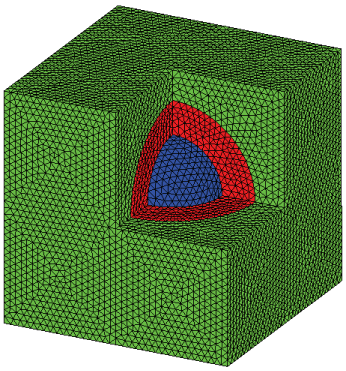

Figure 7: Finite element mesh for modeling the interphase in continuum mechanics.

the hyperparameters of the prior stochastic model of the non-Gaussian matrix-valued random field $\mathbf{x} \mapsto [\mathbb{A}^{\text{meso}}(\mathbf{x})]$ defined in the interphase. For that, the observed random quantity is chosen as the random apparent elasticity matrix $[\mathbb{A}^{\text{app}}]$ related to the domain.The probability density function of $[\mathbb{A}^{\text{app}}]$ is estimated with 200 realizations that are computed by stochastic homogenization, using the finite element solutions of the 6 stochastic BVP with the prior stochastic model $\mathbf{x} \mapsto [\mathbb{A}^{\text{meso}}(\mathbf{x})]$ in the interphase. The likelihood function is evaluated for $[\mathbb{A}^{\text{app, MD}}]$ computed with the the molecular dynamics experiments. About 2 days of computation are required for computing one realization for each configuration, and one month of computation with 8 cores is required for the total computation.

Result of the identification. The optimal values of the hyperparameters are the following: 0.2 for the dispersion parameter, $e/4 = 5 \times 10^{-10}\, m$ for the radial spatial correlation length, and $3.5 \times 10^{-9}\, m$ for the (average) angular spatial correlation length. Figure 8 displays a realization of the random field $\mathbf{x} \mapsto [\mathbb{A}^{\text{meso}}(\mathbf{x})]_{11}$ in the interphase (in GPa) yielding a mean value in the interphase of $5.4\, GPa$ while $[\mathbb{A}^{\text{meso}}]$ is constant in the inclusion and in matrix, and such that $[\mathbb{A}^{\text{meso}}]_{11} = 72\, GPa$ in the inclusion, and $[\mathbb{A}^{\text{meso}}]_{11} = 5\, GPa$ in the matrix.

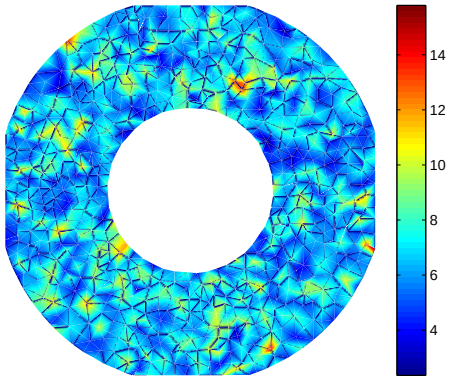

Figure 8: Realization of the random field $\mathbf{x} \mapsto [\mathbb{A}^{\text{meso}}(\mathbf{x})]_{11}$ in the interphase (in GPa).

7 APPLICATION 3

This application is devoted to the multiscale identification [61, 63] of the prior stochastic model of the apparent material properties at mesoscale, using a multiscale experimental digital image correlation [61, 62], at macroscale and at mesoscale.

7.1 Difficulties and multiscale identification

The problem consists of the experimental identification of the vector-valued hyperparameter $\mathbf{b}$ of the prior stochastic model of the non-Gaussian matrix-valued random field $\{[\mathbb{A}^{\text{meso}}(\mathbf{x};\mathbf{b})], \mathbf{x} \in \Omega^{\text{meso}}\}$ that models the apparent elasticity field at mesoscale. The hyperparameter $\mathbf{b}$ is made up of the statistical mean matrix, $E\{[\mathbb{A}^{\text{meso}}(\mathbf{x})]\}$, and other parameters that control the statistical fluctuations such as the spatial-correlation lengths and a dispersion coefficient. The difficulties of this problem are induced by the fact that the mean value $E\{[\mathbb{A}^{\text{meso}}(\mathbf{x})]\}$ cannot directly be identified using only the measurements of the displacement field $\mathbf{u}^{\text{meso}}_{\text{exp}}$ at mesoscale in Ω^{meso}, and requires macroscale measurements. Consequently, some experimental multiscale measurements are required and must be made simultaneously at macroscale and at mesoscale.

7.2 Experimental digital image correlation at macroscale and at mesoscale

Only a single specimen, submitted to a given load applied at macroscale, is tested (see Figure 9). A measurement of the strain field at macroscale is carried out in Ω^{macro} (spatial resolution $10^{-3}\,m$, for instance). Simultaneously, the measurement of the strain field at mesoscale is carried out in Ω^{meso} (spatial resolution $10^{-5}\,m$, for instance).

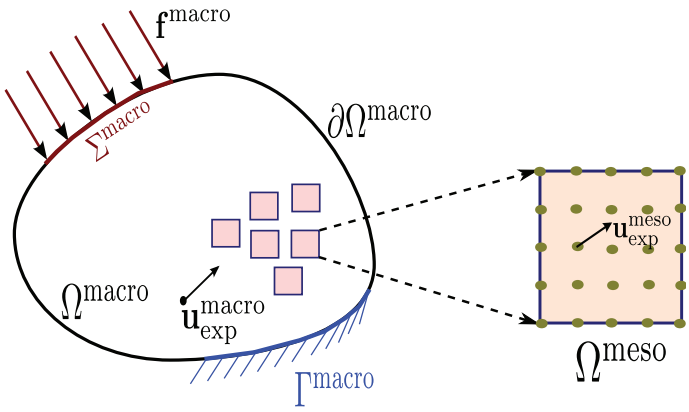

Figure 9: Macroscale and mesoscale measurements of the strain field of a single specimen.

7.3 Hypotheses and strategy for solving the statistical inverse problem

The hypotheses used for solving the statistical inverse problem are the following: (i) there is a separation of macroscale Ω^{macro} from mesoscale Ω^{meso} that is thus a representative volume element (RVE); (ii) at macroscale, the elasticity tensor is constant (independent of $\mathbf{x}$); (iii) at mesoscale, the random apparent elasticity field is homogeneous.
The strategy for solving the statistical inverse problem consists in constructing (i) a prior deterministic model of the macro elasticity matrix, $[\mathbb{A}^{\text{macro}}(\mathbf{a})]$, at macroscale, depending on a vector-

valued parameter $\mathbf{a}$ belonging to an admissible set $\mathcal{A}^{\text{macro}}$, and (ii) a prior stochastic model of the non-Gaussian matrix-valued random field $\{[\mathbb{A}^{\text{meso}}(\mathbf{x};\mathbf{b})], \mathbf{x} \in \Omega^{\text{meso}}\}$ that represents the apparent elasticity field at mesoscale and that depends on a vector-valued hyperparameter $\mathbf{b}$ belonging to an admissible set $\mathcal{B}^{\text{meso}}$.

7.4 Defining numerical indicators for solving the multiscale statistical inverse problem

For solving the multiscale statistical inverse problem, three numerical indicators are introduced, which are:

- a macroscopic numerical indicator $\mathcal{J}_1(\mathbf{a})$ that allows for minimizing the distance between the experimental strain deformation at macroscale and the computed strain deformation at macroscale.
- a mesoscopic numerical indicator $\mathcal{J}_2(\mathbf{b})$ that allows for minimizing the distance between the experimental statistical fluctuations of the strain deformation at mesoscale and the statistical fluctuations of the computed random strain deformation at mesoscale.
- a macroscopic-mesoscopic numerical indicator $\mathcal{J}_3(\mathbf{a},\mathbf{b})$ that allows for minimizing the distance between the elasticity matrix $[\mathbb{A}^{\text{macro}}(\mathbf{a})]$ at macroscale and the effective elasticity matrix $[\mathbb{A}^{\text{eff}}(\mathbf{b})]$ constructed by a stochastic homogenization using the RVE Ω^{meso}.

Macroscopic numerical indicator. The macroscopic numerical indicator $\mathcal{J}_1(\mathbf{a})$ that allows for minimizing the distance between the experimental strain deformation at macroscale and the computed strain deformation at macroscale, is defined by

$$\mathcal{J}_1(\mathbf{a}) = \int_{\Omega^{\text{macro}}} \|\varepsilon^{\text{macro}}_{\text{exp}}(\mathbf{x}) - \varepsilon^{\text{macro}}(\mathbf{x};\mathbf{a})\|^2_F \, d\mathbf{x}\,, \tag{23}$$

in which $\varepsilon^{\text{macro}}_{\text{exp}}(\mathbf{x})$ is the strain field measured in Ω^{macro} and where $\varepsilon^{\text{macro}}(\mathbf{x};\mathbf{a})$ is computed in solving the following boundary value problem at macroscale,

$$\begin{aligned} -\text{div}\,\sigma^{\text{macro}} &= 0 \quad \text{in} \quad \Omega^{\text{macro}}\,, \\ \sigma^{\text{macro}}\mathbf{n}^{\text{macro}} &= \mathbf{f}^{\text{macro}} \quad \text{on} \quad \Sigma^{\text{macro}}\,, \\ \mathbf{u}^{\text{macro}} &= 0 \quad \text{on} \quad \Gamma^{\text{macro}}\,, \\ \sigma^{\text{macro}} = \mathbb{C}^{\text{macro}}(\mathbf{a}) : \varepsilon^{\text{macro}} &\quad , \quad \mathbf{a} \in \mathcal{A}^{\text{macro}}\,, \end{aligned}$$

in which $\mathbb{C}^{\text{macro}}(\mathbf{a})$ is the fourth-order elasticity tensor whose matrix representation is the (6×6) symmetric matrix $[\mathbb{A}^{\text{macro}}(\mathbf{a})]$, and where the prior deterministic model of the macro elasticity matrix, $[\mathbb{A}^{\text{macro}}(\mathbf{a})]$, at macroscale, depends on vector-valued parameter $\mathbf{a} \in \mathcal{A}^{\text{macro}}$.

Mesoscopic numerical indicator. The mesoscopic numerical indicator $\mathcal{J}_2(\mathbf{b})$ allows for minimizing the distance between (i) the normalized dispersion coefficient, $\delta^{\text{meso}}(\mathbf{x};\mathbf{b})$, which characterizes the statistical fluctuations of the computed random strain deformation at mesoscale, and (ii) the corresponding normalized dispersion coefficient, $\delta^{\text{meso}}_{\text{exp}}$, deduced from the experimental strain deformation at mesoscale. This numerical indicator is written as,

$$\mathcal{J}_2(\mathbf{b}) = \int_{\Omega^{\text{meso}}} (\delta^{\text{meso}}_{\text{exp}} - \delta^{\text{meso}}(\mathbf{x};\mathbf{b}))^2 \, d\mathbf{x}\,. \tag{24}$$

The quantities $\delta^{\text{meso}}(\mathbf{x};\mathbf{b})$ and $\delta^{\text{meso}}_{\text{exp}}$ are defined by

$$\delta^{\text{meso}}(\mathbf{x};\mathbf{b}) = \frac{\sqrt{V^{\text{meso}}(\mathbf{x};\mathbf{b})}}{\|\underline{\varepsilon}^{\text{meso}}_{\text{exp}}\|_F} \quad , \quad V^{\text{meso}}(\mathbf{x};\mathbf{b}) = E\{\|\varepsilon^{\text{meso}}(\mathbf{x};\mathbf{b}) - \underline{\varepsilon}^{\text{meso}}(\mathbf{b})\|_F^2\}$$

$$\delta^{\text{meso}}_{\text{exp}} = \frac{\sqrt{V^{\text{meso}}_{\text{exp}}}}{\|\underline{\varepsilon}^{\text{meso}}_{\text{exp}}\|_F} \quad , \quad V^{\text{meso}}_{\text{exp}} = \frac{1}{|\Omega^{\text{meso}}|}\int_{\Omega^{\text{meso}}} \|\varepsilon^{\text{meso}}_{\text{exp}}(\mathbf{x}) - \underline{\varepsilon}^{\text{meso}}_{\text{exp}}\|_F^2 \, d\mathbf{x}$$

in which $\|\cdot\|_F$ is the Frobenius norm, and where the spatial averagings $\underline{\varepsilon}^{\text{meso}}(\mathbf{b})$ and $\underline{\varepsilon}^{\text{meso}}_{\text{exp}}$ on the RVE, Ω^{meso}, are defined by

$$\underline{\varepsilon}^{\text{meso}}(\mathbf{b}) = \frac{1}{|\Omega^{\text{meso}}|}\int_{\Omega^{\text{meso}}} \varepsilon^{\text{meso}}(\mathbf{x};\mathbf{b})\, d\mathbf{x} \quad , \quad \underline{\varepsilon}^{\text{meso}}_{\text{exp}} = \frac{1}{|\Omega^{\text{meso}}|}\int_{\Omega^{\text{meso}}} \varepsilon^{\text{meso}}_{\text{exp}}(\mathbf{x})\, d\mathbf{x}\,.$$

For all $\mathbf{b}$ fixed in $\mathcal{B}^{\text{meso}}$, the mesoscale strain field $\varepsilon^{\text{meso}}(\mathbf{x};\mathbf{b})$ is calculated by solving the mesoscale boundary value problem,

$$\begin{aligned} -\text{div}\,\sigma^{\text{meso}} &= 0 \quad \text{in} \quad \Omega^{\text{meso}}\,, \\ \mathbf{U}^{\text{meso}} &= \mathbf{u}^{\text{meso}}_{\text{exp}} \quad \text{on} \quad \partial\Omega^{\text{meso}}\,, \\ \sigma^{\text{meso}} &= \mathbb{C}^{\text{meso}}(\mathbf{b}) : \varepsilon^{\text{meso}} \quad , \quad \mathbf{b} \in \mathcal{B}^{\text{meso}}\,, \end{aligned}$$

in which $\mathbf{u}^{\text{meso}}_{\text{exp}}$ is the displacement field measured on the boundary $\partial\Omega^{\text{meso}}$ of the mesoscale domain Ω^{meso}, where $\mathbb{C}^{\text{meso}}(\mathbf{b})$ is the non-Gaussian fourth-order tensor-valued elasticity field whose matrix representation is the prior stochastic model of the non-Gaussian matrix-valued random field $\{[\mathbb{A}^{\text{meso}}(\mathbf{x};\mathbf{b})], \mathbf{x} \in \Omega^{\text{meso}}\}$ (random apparent elasticity field at mesoscale) that depends on vector-valued hyperparameter $\mathbf{b} \in \mathcal{B}^{\text{meso}}$. Since Ω^{meso} is assumed to be an RVE, then for all $\mathbf{b}$ in $\mathcal{B}^{\text{meso}}$, we have $\underline{\varepsilon}^{\text{meso}}(\mathbf{b}) = \underline{\varepsilon}^{\text{meso}}_{\text{exp}}$ almost surely (a.s).

Macroscopic-mesoscopic numerical indicator. The macroscopic-mesoscopic numerical indicator $\mathcal{J}_3(\mathbf{a},\mathbf{b})$ allows for minimizing the distance between the macro elasticity tensor $\mathbb{C}^{\text{macro}}(\mathbf{a})$ at macroscale and the effective elasticity matrix $[\mathbb{A}^{\text{eff}}(\mathbf{b})]$ constructed by a stochastic homogenization using the RVE Ω^{meso}. This indicator is written as,

$$\mathcal{J}_3(\mathbf{a},\mathbf{b}) = \|[\mathbb{A}^{\text{macro}}(\mathbf{a})] - E\{[\mathbb{A}^{\text{eff}}(\mathbf{b})]\}\|_F^2\,. \tag{25}$$

The stochastic homogenization (from meso to macro) is formulated in homogeneous constraints (that is better adapted for the 2D plane stresses) with $\sigma^{\text{meso}} = \mathbb{C}^{\text{meso}}(\mathbf{b}) : \varepsilon^{\text{meso}}$.

7.5 Statistical inverse problem formulated as a multi-objective optimization problem

The statistical inverse problem then consists in identifying the optimal values $\mathbf{a}^{\text{macro}}$ of the vector-valued parameter $\mathbf{a}$ in $\mathcal{A}^{\text{macro}}$ and $\mathbf{b}^{\text{meso}}$ of the vector-valued hyperparameter $\mathbf{b}$ in $\mathcal{B}^{\text{meso}}$ by solving the following multi-objective optimization problem,

$$(\mathbf{a}^{\text{macro}}, \mathbf{b}^{\text{meso}}) = \arg \min_{\mathbf{a}\in\mathcal{A}^{\text{macro}}, \mathbf{b}\in\mathcal{B}^{\text{meso}}} \boldsymbol{\mathcal{J}}(\mathbf{a},\mathbf{b})\,, \tag{26}$$

in which $\min \boldsymbol{\mathcal{J}}(\mathbf{a},\mathbf{b}) = (\min \mathcal{J}_1(\mathbf{a}), \min \mathcal{J}_2(\mathbf{b}), \min \mathcal{J}_3(\mathbf{a},\mathbf{b}))$. For solving this multi-objective optimization problem (see [61, 63]):

- the deterministic BVP at macroscale is discretized using the FEM;
- the stochastic BVP at mesoscale is discretized using the FEM and is solved using the Monte Carlo method;
- the multi-objective optimization problem is solved using the genetic algorithm, and a Pareto front is iteratively constructed at each generation of the genetic algorithm;
- the initial value $\mathbf{a}^{(0)}$ of $\mathbf{a} \in \mathcal{A}^{\text{macro}}$ is computed by solving the optimization problem, $\mathbf{a}^{(0)} = \arg\min_{\mathbf{a}\in\mathcal{A}^{\text{macro}}} \mathcal{J}_1(\mathbf{a})$, using the simplex algorithm;
- the hyperparameter $\mathbf{b}$ is chosen in $\mathcal{B}^{\text{meso}}$ as the point on the Pareto front that minimizes the distance between the Pareto front and the origin.

7.6 Multiscale experimental measurements of a cortical bone in 2D plane stresses

The experimental measurements for the identification of the elasticity field at mesoscale of a heterogeneous microstructure by multiscale digital image correlation [61, 63] have been carried out at LMS of Ecole Polytechnique [61, 62] (see Figure 10). The specimen is a cubic bovine cortical bone with dimensions $0.01 \times 0.01 \times 0.01\, m^3$. The dimensions, the spatial resolution and the applied force to the specimen for the multiscale measurements (see Figure 11) are the following:

- $\Omega^{\text{macro}} : 0.01 \times 0.01\, m^2$ meshed with a 10×10-points grid yielding a spatial resolution of $10^{-3} \times 10^{-3}\, m^2$;
- $\Omega^{\text{meso}} : 0.001 \times 0.001\, m^2$ meshed with a 100×100-points grid yielding a spatial resolution of $10^{-5} \times 10^{-5}\, m^2$.
- Applied force: $9,000\, N$.

A comparison between a reference image and a deformed image obtained by digital image correlation experimental method is shown in Figure 12 for the cubic bovine cortical bone sample at macroscale. The experimental displacement field measured at macroscale is shown in Figure 13, and the experimental displacement field measured at mesoscale is shown in Figure 14.

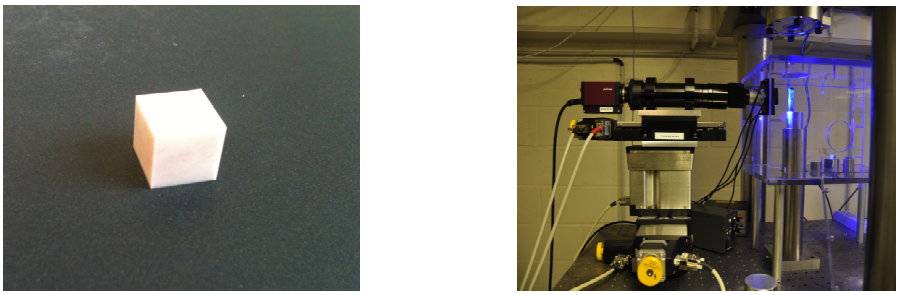

Figure 10: Specimen of the cubic bovine cortical bone (left) and measuring bench (right).

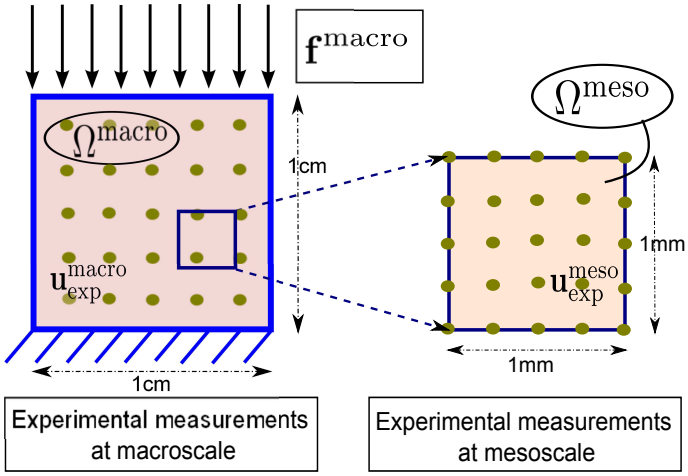

Figure 11: Dimensions, applied force and boundary conditions of the specimen for the multiscale measurements.

Figure 12: Comparison between a reference image (left) and a deformed image (right) at macroscale for the cubic bovine cortical bone sample.

7.7 Stochastic computational model and results obtained by the multiscale identification procedure

Stochastic computational model. The details concerning the construction of the stochastic computational model can be found in [61, 63]. A 2D-plane-stresses modeling is used. At macroscale, the material is assumed to be homogeneous, transverse isotropic, linear elastic, and the parameter $\mathbf{a}$ is defined as $\mathbf{a} = (E_T^{\text{macro}}, \nu_T^{\text{macro}})$ in which E_T^{macro} is the transverse Young modulus and ν_T^{macro} is the Poisson coefficient. At mesoscale, the material is assumed to be heterogeneous, anisotropic, linear elastic, the statistical mean value is assumed to be transverse isotropic, and the statistical fluctuations are anisotropic. The prior stochastic model of the apparent elasticity field at mesoscale is deduced from the full anisotropic stochastic case previously described and coincides with the prior stochastic model introduced in [49, 40, 44]. The hyperparameter $\mathbf{b}$ is defined by $\mathbf{b} = (\underline{E}_T, \underline{\nu}_T, L, \delta)$ in which $\underline{E}_T$ is the statistical mean value of the transverse Young modulus, $\underline{\nu}_T$ is the statistical mean value of the Poisson coefficient, L is the spatial-correlation length that is assumed equal for the three cartesian directions, and where δ is the dispersion parameter that allows the statistical fluctuations of the apparent elasticity field to be controlled.

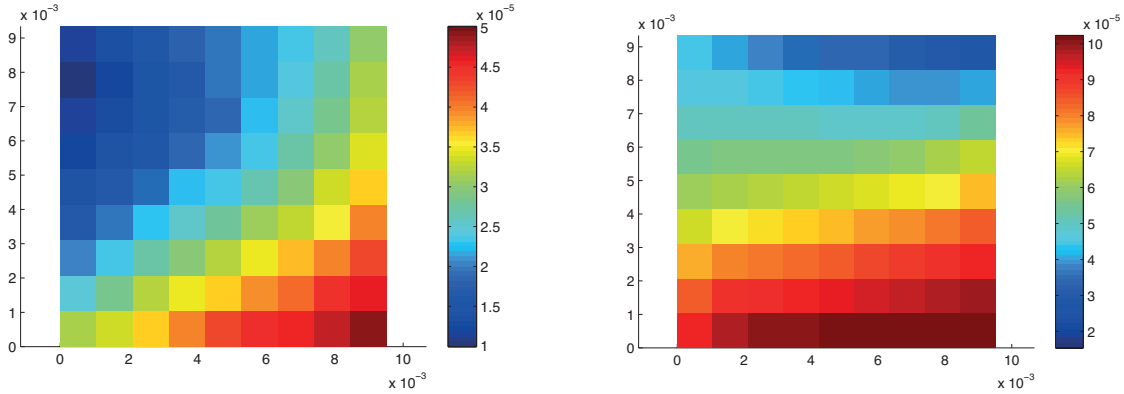

Figure 13: Component $\{\mathbf{u}_{\exp}^{\text{macro}}\}_1$ in direction x_1 (left figure) and component $\{\mathbf{u}_{\exp}^{\text{macro}}\}_2$ in direction x_2 (right) for the experimental displacement at macroscale.

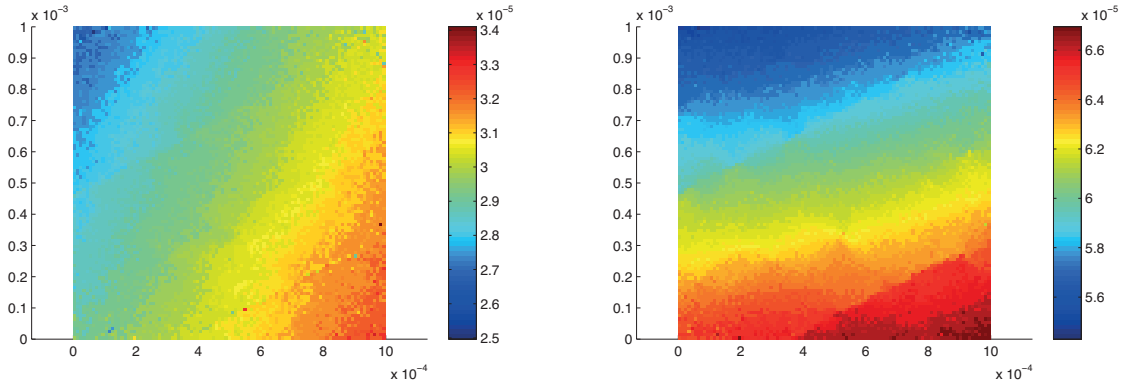

Figure 14: Component $\{\mathbf{u}_{\exp}^{\text{meso}}\}_1$ in direction x_1 (left figure) and component $\{\mathbf{u}_{\exp}^{\text{meso}}\}_2$ in direction x_2 (right figure) for the experimental displacement at mesoscale.

Results obtained by the multiscale identification procedure[61, 63]. The optimal value $\mathbf{a}^{\text{macro}} = (E_T^{\text{macro}}, \nu_T^{\text{macro}})$ is such that $E_T^{\text{macro}} = 6.74 \times 10^9\, Pa$ and $\nu_T^{\text{macro}} = 0.32$. The optimal values of the components of $\mathbf{b}^{\text{meso}} = (\underline{E}_T^{\text{meso}}, \underline{\nu}_T^{\text{meso}}, L^{\text{meso}}, \delta^{\text{meso}})$ are $\underline{E}_T^{\text{meso}} = 6.96 \times 10^9\, Pa$, $\underline{\nu}_T^{\text{meso}} = 0.37$, $L^{\text{meso}} = 5.06 \times 10^{-5}\, m$, and $\delta^{\text{meso}} = 0.28$. The identified spatial-correlation length is in agreement with the assumption introduced concerning the separation of the macroscopic and the mesoscopic scales, and is of the same order of magnitude than the distance between adjacent lamellae or osteons in bovine cortical femur. The identified values of **a** and **b** are coherent with the values published in the literature.

8 CONCLUSION

A general methodology and some algorithms have been presented for identifying a non-Gaussian tensor-valued random field in high stochastic dimension. This random field can constitute the parameter of a boundary value problem (BVP) for which only partial and limited experimental data are available for the observation. Taking into account the high dimensionality and the non-Gaussian character of the random field that has to be identified by solving a statistical inverse problem, an algebraic prior stochastic modeling and/or an adapted stochastic representation must be constructed so that the inverse problem can be effectively solved. Three applications in different domains have been presented and demonstrates the efficiency of the

methodology proposed.

REFERENCES

[1] J. Kaipio, E. Somersalo E, *Statistical ans Computational Inverse Problems*. Springer-Verlag, New York, 2005.

[2] A. Tarantola, *Inverse problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia, 2005.

[3] R.J. Serfling, *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York, 1980.

[4] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework. *Journal of Engineering Mechanics-ASCE*, **124**(4), 455–461, 1998.

[5] J.M. Bernardo, A.F.M. Smith, *Bayesian Theory*. John Wiley & Sons, Chichester, 2000.

[6] J.C. Spall, *Introduction to Stochastic Search and Optimization*. John Wiley & Sons, Hoboken, 2003.

[7] Congdon P (2007) Bayesian Statistical Modelling. Second Edition, John Wiley & Sons, Chichester.

[8] B.P. Carlin, T.A. Louis, *Bayesian Methods for Data Analysis*. Third Edition, Chapman & Hall / CRC Press, Boca Raton, 2009.

[9] A.M. Stuart, Inverse problems: A bayesian perspective. *Acta Numerica* **19**, 451–559, 2010.

[10] R. Ghanem, P.D. Spanos, *Stochastic Finite Elements: a Spectral Approach*. Springer-Verlag, New York, 1991.

[11] R. Ghanem, R.M. Kruger, Numerical solution of spectral stochastic finite element systems. *Computer Methods in Applied Mechanics and Engineering*, **129**(3), 289–303, 1996.

[12] R. Ghanem, P.D. Spanos, *Stochastic Finite Elements: A spectral Approach (revised edition)*. Dover Publications, New York, 2003.

[13] R. Ghanem, S. Dham, Stochastic finite element analysis for multiphase flow in heterogeneous porous media. *Transport in Porous Media*, **32**, 239–262, 1998.

[14] B. Puig, F. Poirion, C. Soize, Non-Gaussian simulation using Hermite polynomial expansion: Convergences and algorithms. *Probabilistic Engineering mechanics*, **17**(3), 253–264, 2002.

[15] B.J. Debusschere, H.N. Najm, P.P. Pebay, O.M. Knio, R. Ghanem, O. Le Maître O, Numerical challenges in the use of polynomial chaos representations for stochastic processes. *SIAM Journal on Scientific Computing*, **26**(2), 698–719, 2004.

[16] O.M. Knio, O.P. Le Maître, Uncertainty propagation in CFD using polynomial chaos decomposition. *Fluid Dynamics Reserach*, **38**(9), 616–640, 2006.

[17] A. Doostan, R. Ghanem, J. Red-Horse, Stochastic model reduction for chaos representations. *Computer Methods in Applied Mechanics and Engineering*, **196**(37-40), 3951–3966, 2007.

[18] D. Ghosh, R. Ghanem, Stochastic convergence acceleration through basis enrichment of polynomial chaos expansions. *International Journal for Numerical Methods in Engineering*, **73**(2), 162–184, 2008.

[19] H.H. Najm, Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. *Annual Review of Fluid Mechanics*, **41**, 35–52, 2009.

[20] O.P. Le Maître, O.M. Knio, *Spectral Methods for Uncertainty Quantification with Applications to Computational Fluid Dynamics*. Springer, Heidelberg, 2010.

[21] A. Nouy, Proper Generalized Decomposition and separated representations for the numerical solution of high dimensional stochastic problems. *Archives of Computational Methods in Engineering*, **16**(3), 403–434, 2010.

[22] C. Soize, R. Ghanem, Physical systems with random uncertainties : Chaos representation with arbitrary probability measure. *SIAM Journal on Scientific Computing*, **26**(2), 395–410, 2004.

[23] O.P. Le Maitre, O.M. Knio OM, H.N. Najm, Uncertainty propagation using Wiener-Haar expansions. *Journal of Computational Physics*, **197**(1), 28–57, 2004.

[24] D. Lucor, C.H. Su, G.E. Karniadakis, Generalized polynomial chaos and random oscillators. *International Journal for Numerical Methods in Engineering*, **60**(3), 571–596, 2004.

[25] X.L. Wan XL, G.E. Karniadakis, Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM Journal on Scientific Computing*, **28**(3), 901–928, 2006.

[26] O.G. Ernst, A. Mugler, H.J. Starkloff, E. Ullmann, On the convergence of generalized polynomial chaos expansions. *ESAIM: Mathematical Modelling and Numerical Analysis*, **46**(2), 317–339, 2012.

[27] C. Soize, C. Desceliers, Computational aspects for constructing realizations of polynomial chaos in high dimension. *SIAM Journal On Scientific Computing*, **32**(5), 2820–2831, 2010.

[28] G. Perrin, C. Soize, D. Duhamel, C. Funfschilling, Identification of polynomial chaos representations in high dimension from a set of realizations. *SIAM Journal on Scientific Computing*, **34**(6), A2917–A2945, 2012.

[29] C. Soize, R. Ghanem, Reduced chaos decomposition with random coefficients of vector-valued random variables and random fields. *Computer Methods in Applied Mechanics and Engineering*, **198**(21-26), 1926–1934, 2009.

[30] R. Tipireddy, R. Ghanem, Basis adaptation in homogeneous chaos spaces. *Journal of Computational Physics*, **259**, 304–317, 2014.

[31] C. Soize, Polynomial chaos expansion of a multimodal random vector. *SIAM/ASA Journal on Uncertainty Quantification*, **3**(1), 34-60, 2015.

[32] C. Desceliers, R. Ghanem, C. Soize, Maximum likelihood estimation of stochastic chaos representations from experimental data. *International Journal for Numerical Methods in Engineering*, **66**(6), 978–1001, 2006.

[33] C. Desceliers, C. Soize, R. Ghanem, Identification of chaos representations of elastic properties of random media using experimental vibration tests. *Computational Mechanics*, **39**(6), 831–838, 2007.

[34] J. Guilleminot, C. Soize, D. Kondo, C. Binetruy, Theoretical framework and experimental procedure for modelling volume fraction stochastic fluctuations in fiber reinforced composites. *International Journal of Solid and Structures*, **45**(21), 5567–5583, 2008.

[35] J. Guilleminot, C. Soize, D. Kondo, Mesoscale probabilistic models for the elasticity tensor of fiber reinforced composites: experimental identification and numerical aspects. *Mechanics of Materials*, **41**(12), 1309–1322, 2009.

[36] S. Das, R. Ghanem, S. Finette, Polynomial chaos representation of spatio-temporal random field from experimental measurements. *Journal of Computational Physics*, **228**, 8726–8751, 2009.

[37] S. Das, R. Ghanem, J.C. Spall, Asymptotic sampling distribution for polynomial chaos representation from data: a maximum entropy and fisher information approach. *SIAM Journal on Scientific Computing*, **30**(5), 2207–2234, 2008.

[38] R. Ghanem, R. Doostan, J. Red-Horse, A probability construction of model validation. *Computer Methods in Applied Mechanics and Engineering*, **197**(29-32), 2585–2595, 2008.

[39] C. Soize, Identification of high-dimension polynomial chaos expansions with random coefficients for non-Gaussian tensor-valued random fields using partial and limited experimental data. *Computer Methods in Applied Mechanics and Engineering*, **199**(33–36), 2150–2164, 2010.

[40] C. Soize, *Stochastic models of uncertainties in computational mechanics*. American Society of Civil Engineers (ASCE), Reston, 2012.

[41] M. Arnst, R. Ghanem, C. Soize, Identification of Bayesian posteriors for coefficients of chaos expansions. *Journal of Computational Physics*, **229**(9), 3134–3154, 2010.

[42] C. Soize, A computational inverse method for identification of non-Gaussian random fields using the Bayesian approach in very high dimension. *Computer Methods in Applied Mechanics and Engeneering*, **200**(45-46), 3083–3099, 2011.

[43] A. Nouy, C. Soize, Random fields representations for stochastic elliptic boundary value problems and statistical inverse problems. *European Journal of Applied Mathematics*, **25**(3), 339–373, 2014.

[44] C. Soize, Random vectors and random fields in high dimension. Parametric model-based representation, identification from data, and inverse problems, pp. 1–65, in *Handbook of Uncertainty Quantification*, edited by R. Ghanem, H. Owhadi, and D. Higdon, Springer-Verlag (to appear 2015).

[45] C. Soize, Random matrix models and nonparametric method for uncertainty quantification, pp. 1–84, in *Handbook of Uncertainty Quantification*, edited by R. Ghanem, H. Owhadi, and D. Higdon, Springer-Verlag (to appear 2015).

[46] G. Perrin, C. Soize, D. Duhamel, C. Funfschilling, Karhunen-Loève expansion revisited for vector-valued random fields: scaling, errors and optimal basis. *Journal of Computational Physics*, **242**(1), 607–622, 2013.

[47] G. Perrin, C. Soize, D. Duhamel, C. Funfschilling, A posterior error and optimal reduced basis for stochastic processes defined by a set of realizations. *SIAM/ASA Journal of Uncertainty Quantification*,**2**, 745–762, 2014.

[48] C. Soize, Random-field model for the elasticity tensor of anisotropic random media. *Compte Rendu Mecanique*, **332**, 1007–1012, 2004.

[49] C. Soize, Non Gaussian positive-definite matrix-valued random fields for elliptic stochastic partial differential operators. *Computer Methods in Applied Mechanics and Engineering*, **195**(1-3), 26–64, 2006.

[50] C. Soize, Construction of probability distributions in high dimension using the maximum entropy principle. Applications to stochastic processes, random fields and random matrices. *International Journal for Numerical Methods in Engineering*, **76**(10), 1583–1611, (2008).

[51] J. Guilleminot, A. Noshadravan, C. Soize, R. Ghanem, A probabilistic model for bounded elasticity tensor random fields with application to polycrystalline microstructures. *Computer Methods in Applied Mechanics and Engineering*, **200**, 1637-1648, 2011.

[52] J. Guilleminot, C. Soize, Non-Gaussian positive-definite matrix-valued random fields with constrained eigenvalues: application to random elasticity tensors with uncertain material symmetries. *International Journal for Numerical Methods in Engineering*, **88**(11), 1128–1151, 2011.

[53] J. Guilleminot, C. Soize, Probabilistic modeling of apparent tensors in elastostatics: a MaxEnt approach under material symmetry and stochastic boundedness constraints. *Probabilistic Engineering Mechanics*, **28**, 118–124, 2012.

[54] J. Guilleminot, C. Soize; R. Ghanem, Stochastic representation for anisotropic permeability tensor random fields. *International Journal for Numerical and Analytical Methods in Geomechanics*, **36**(13), 1592–1608, 2012.

[55] J. Guilleminot, C. Soize, Generalized stochastic approach for constitutive equation in linear elasticity: A random matrix model. *International Journal for Numerical Methods in Engineering*, **90**(5), 613–635, 2012.

[56] J. Guilleminot, C. Soize, Stochastic model and generator for random fields with symmetry properties: application to the mesoscopic modeling of elastic random media. *Multiscale Modeling and Simulation (A SIAM Interdisciplinary Journal)*, **11**(3), 840–870, 2013.

[57] J. Guilleminot, T.T. Le, C. Soize, Stochastic framework for modeling the linear apparent behavior of complex materials: application to random porous materials with interphases. *Acta Mechanica Sinica*, **29**(6), 773–782, 2013.

[58] J. Guilleminot, C. Soize, ISDE-based generator for a class of non-gaussian vector-valued random fields in uncertainty quantification. *SIAM Journal on Scientific Computing*, **36**(6), A2763-A2786, 2014.

[59] G. Perrin, C. Soize, D. Duhamel, C. Funfschilling, Track irregularities stochastic modeling. *Probabilistic Engineering Mechanics*, **34**, 123–130, 2013.

[60] T.T. Le, J. Guilleminot, C. Soize, Stochastic continuum modeling of random interphases from atomistic simulations. Application to a polymer nanocomposite. Submitted in 2015.

[61] M. T. Nguyen, *Identification multi-échelle du champ délasticité apparent stochastique de microstructures hétérogènes. Application à un tissu biologique*. Doctoral Thesis, Université Paris-Est, Marne-la-Vallée, France, October 8, 2013.

[62] M. T. Nguyen, J. M. Allain, H. Gharbi, C. Desceliers, C. Soize, Experimental measurements for identification of the elasticity field at mesoscale of a heterogeneous microstructure by multiscale digital image correlation, submitted 2015.

[63] M. T. Nguyen, C. Desceliers, C. Soize, J. M. Allain, H. Gharbi, Multiscale identification of the random elasticity field at mesoscale of a heterogeneous microstructure using multiscale experimental observations. *International Journal for Multiscale Computational Engineering*, accepted for publication 12 September 2014.

[64] P.-A. Absil, R. Mahony, R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, New Jersey, 2008.

[65] A. Ta, D. Clouteau, R. Cottereau, Modeling of random anisotropic elastic media and impact on wave propagation. *European Journal of Computational Mechanics*, **19**, 241-253, 2010.

[66] G. Perrin, Random fields and associated statistical inverse problems for uncertainty quantification. Application to railway track geometries for high-speed trains dynamical responses and risk assessment, Doctoral Thesis, Université Paris-Est, Marne-la-Vallée, France, September 24, 2013.